Modelling Life Tables with Advanced Ages: an Extreme Value Theory Approach [1]

Fei Huang
(Joint Work with Ross Maller and Xu Ning)

Senior Lecturer in Actuarial Studies,
School of Risk and Actuarial Studies
UNSW Business School

28th Colloquium on Pensions and Retirement Research
8 December, 2020

[1] This paper has been accepted in *Insurance: Mathematics and Economics*. Download the paper here

The number of Aussies aged over 100 grew by more than 30 per cent between 2013 and 2018.

Contributions

- We propose new mortality models which can incorporate advanced ages:
  1. Smooth Threshold Life Table (STLT) Model
  2. Dynamic Smooth Threshold Life Table (DSTLT) Model
- We discover a new law of mortality, called "advanced-age mortality acceleration".

Data

HMD Dataset

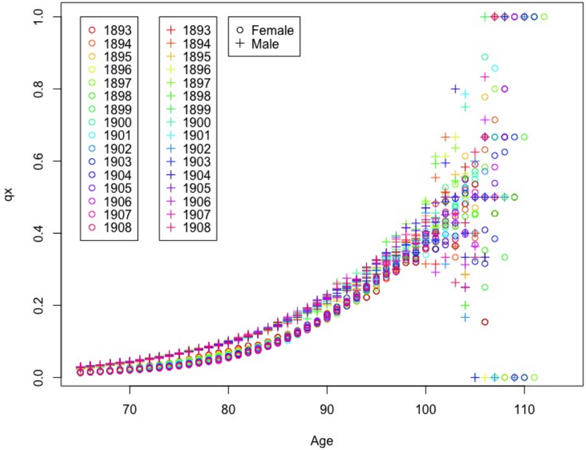
1. For ages between 65 to 100+, we use the interval-censored data from Human Mortaltiy Database (HMD) for the Netherlands.

Augmented Dataset (HMD+CBS)

1. For ages 65 - 92, we use the interval-censored cohort data from Human Mortaltiy Database (HMD) for the Netherlands.
2. For ages 92 and above, we use individual-level ages at death data from the Centraal Bureau voor de Statistiek (CBS) of the Netherlands, 1986-2015.

Augmented Data: HMD+CBS

We plot the $q_x = d_x / l_x$: the empirical conditional probability of death between ages x and $x + 1$ for individuals surviving to age x, of the augmented dataset (HMD+CBS).



Notations

- X: age at death of an individual from the population, a continuous random variable
- $f(x)$: probability density function of X
- $F(x)$: cumulative distribution function (cdf) of X
- $S(x) = 1 - F(x)$: survival function of F
- $h(x) = f(x)/(1 - F(x))$: force of mortality, or hazard function, corresponding to F
- d_x: the number of deaths between integer ages x and $x + 1$
- E_x: the population exposed to the risk of death between integer ages x and $x + 1$. E_x is approximated by the midyear population at age x
- l_x: the number of survivors to age x
- $m_x = d_x/E_x$: the central rate of death at age x
- $q_x = d_x/l_x$: the empirical conditional probability of death between ages x and $x + 1$ for individuals surviving to age x.

Some Existing Models for Advanced Age Mortality Modelling

Static Models

1. Gompertz-Makeham Law
2. Logistic models
3. Heligman-Pollard model
4. Coale-Kisker method – used by Lee & Carter (1992)
5. Extreme Value Theory based method (Watts et al., 2006, Thatcher, 1999)
   Threshold Life Table (TLT) – Li et al. (2008)

Dynamic Models

1. Combined Models
   e.g. Lee-Carter+Coale-Kisker; Lee-Carter+TLT
2. Watts-Dupuis-Jones (WDJ) Model – Watts et al. (2006)
   Only for highest attained ages
3. Cairns-Blake-Dowd Model – Cairns et al. (2006)
   Gompertz model with time-varying parameters

Gompertz-Makeham Law of Mortality

The force of mortality in the Gompertz model is

$$h(x) = B\exp(Cx). \tag{1}$$

Gompertz modelled adult mortality with two parameters:

- a positive scale parameter B that represents the level of mortality,
- and a positive shape parameter C that measures the rate of increase in mortality with age.

The Makeham, or *Gompertz-Makeham*, law of mortality has an extra parameter A represents mortality resulting from causes.

$$h(x) = A + B\exp(Cx). \tag{2}$$

Logistic Models

A general class of logistic models was formulated by Perks (1932). he hazard function in *Perks's model* is of the form

$$h(x) = A + \frac{B\exp(Cx)}{1 + D\exp(Cx)}. \quad (3)$$

Setting $A = 0$ in (3) gives the three-parameter *Beard model* (Beard, 1971):

$$h(x) = \frac{B\exp(Cx)}{1 + D\exp(Cx)}. \quad (4)$$

Kannisto (1992) noticed that modern data for $h(x)$ at high ages can be well-fitted by one of the simplest forms of the logistic model, in which $\text{logit}(h(x))$ is a linear function of x. The resulting *Kannisto model* specifies the hazard function as

$$h(x) = \frac{B\exp(Cx)}{1 + B\exp(Cx)}. \quad (5)$$

Heligman-Pollard Model

$$\frac{q_x}{1-q_x} = A^{(x+B)^C} + D\exp(-E[\ln x - \ln F]^2) + GH^x, \quad x = 1, 2, 3, \ldots. \tag{6}$$

- A, B, C, D, E, F, G, H are constants
- The terms $A^{(x+B)^C}$, $D\exp(-E[\ln x - \ln F]^2)$ and GH^x represent early childhood mortality, accidental mortality, and senescent mortality, respectively.
- The old age mortality component is similar to the Gompertz-Makeham law of mortality.

Coale-Kisker Model

$$k(x) = k(x-1) - R, \quad x \geq x_0, \tag{7}$$

where $k(x) = \ln(m_x/m_{x+1})$, R is a constant to be determined, and the extrapolation starts at integer age x_0. Applying the formula up to age $x = x_1$, we find

$$R = \frac{(x_1 - x_0)k(x_0) + \ln(m_{x_0}) - \ln(m_{x_1})}{1 + 2 + \cdots + (x_1 - x_0)}. \tag{8}$$

The method requires an assumption of the age x_1 at which the life table is closed as well as the value of the central death rate at this closing age. The age from which to start extrapolating also must be subjectively decided. Coale & Kisker (1990) use $x_0 = 84, x_1 = 110, m_{x_1} = 1.0$.

Cairns-Blake-Dowd Model

$$\log h_{i,j} = \kappa_{0,j} + \kappa_{1,j}(x_i - \bar{x}), \tag{9}$$

where $h_{i,j}$ is the force of mortality in year j at age x_i, and $\kappa_{0,j}, \kappa_{1,j}$, are latent time-varying variables typically modelled as a bivariate random walk with drift. The model can be rewritten in the form of a Gompertz model with time-varying parameters:

$$\log h_{i,j} = \log B_j + x_i \log C_j, \tag{10}$$

where $B_j = \exp(\kappa_{0,j} - \kappa_{1,j}\bar{x})$ and $C_j = \exp(\kappa_{1,j})$.

Our Proposed Models

- Smooth Threshold Life Table (STLT) Model
- Dynamic Smooth Threshold Life Table (DSTLT) Model

The Threshold Life Table (TLT) Model, Li et al. (2008)

The model is piecewise, and comprises the non-tail part of the distribution following the Gompertz-Makeham law of mortality, with the tail following a Generalized Pareto Distribution (GPD). The model is completely defined by:

$$F(x) = 1 - \exp\left(-\frac{B}{\ln C}(C^x - 1)\right) \qquad x \leq N \tag{11}$$

and

$$F(x) = F_\gamma(x) = \begin{cases} 1 - S(N)(1 + \gamma(\frac{x-N}{\theta}))^{-\frac{1}{\gamma}}, & \gamma > 0,\ x > N \\ 1 - S(N)\exp(-(\frac{x-N}{\theta})), & \gamma = 0,\ x > N \\ 1 - S(N)(1 - |\gamma|(\frac{x-N}{\theta}))^{\frac{1}{|\gamma|}}, & \gamma < 0,\ N < x < N + \frac{\theta}{|\gamma|}, \end{cases} \tag{12}$$

which ensures the continuity but not smoothness at the threshold age N, see Li et al. (2008).

Adding a Smooth Constraint at N

We want $h_1(N) = h_2(N)$ where h_1 is the hazard function corresponding to (11) and h_2 the hazard function corresponding to (12). We have

$$\frac{1}{\theta + \gamma(x - N)} = C^x B \qquad \text{at } x = N$$
$$\Rightarrow \theta = \frac{1}{C^N B}.$$

Implications:

1. Removes the discontinuity in the force of mortality.
2. Provides a functional link between the Gompertz and GPD distributions, which makes the extreme age modelling more robust.
3. Satisfies the *compensation law of mortality*

The Smooth TLT (STLT) Model

So the TLT model now becomes the STLT model with the parameter θ eliminated:

$$F(x) = 1 - \exp\left(-\frac{B}{\ln C}(C^x - 1)\right) \qquad x \leq N; \qquad (13)$$

and

$$F(x) = \begin{cases} 1 - S(N)(1 + \gamma(C^N B(x - N)))^{-\frac{1}{\gamma}}, & \gamma > 0,\ x > N \\ 1 - S(N)\exp(-(C^N B(x - N))), & \gamma = 0,\ x > N \\ 1 - S(N)(1 - |\gamma|(C^N B(x - N)))^{\frac{1}{|\gamma|}}, & \gamma < 0,\ N < x < N + \frac{\theta}{|\gamma|}, \end{cases} \qquad (14)$$

Parameters Estimation – MLE

The likelihood, considering interval-censored data, can be written as

$$L(B, C, \gamma; N) = \left[\prod_{x=65}^{N-1} \left(\frac{S(x) - S(x+1)}{S(65)}\right)^{d_x} \prod_{x=N}^{\tau-1} \left(\frac{S(x) - S(x+1)}{S(65)}\right)^{d_x}\right] \left(\frac{S(\tau)}{S(65)}\right)^{l_\tau} \tag{15}$$

where, τ is the observed maximum attained age for the cohort.

$$l_1(B, C; N) = \sum_{x=65}^{N-1} d_x \ln(S(x) - S(x+1)) + l_N \ln(S(N)) - l_{65} \ln(S(65)), \tag{16}$$

where $S(x) = \exp(-B/\ln C(C^x - 1))$, and the second being

$$l_2(\gamma, \theta; N) = \sum_{x=N}^{\tau-1} d_x \ln\left(S(x) - S(x+1)\right) + l_\tau \ln\left(S(\tau)\right) - l_N \ln(S(N)), \tag{17}$$

where $S(x)/S(N) = (1 + \gamma((x - N)/\theta))^{-1/\gamma}$.
$l = l_1 + l_2$.

Model Fitting

We use a constructed hypothetical cohort. The required values of d_x and l_x can be computed using the actual probabilities of death q_x by assuming an arbitrary number, say, 100,000, for the size of radix (l_0).

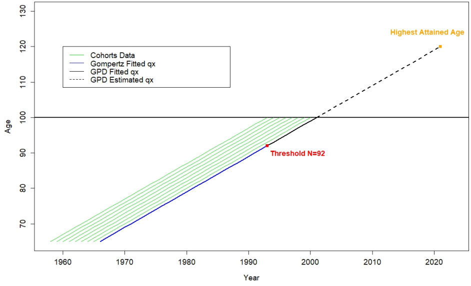


Figure 2: *Schematic for threshold life table fit to the cohort, N=92*

Highest Attained Age

Let $Y = X - N$ denote the exceedance of an individual aged X over threshold age N. In both the TLT and STLT, the distribution of Y conditional on $X > N$ is a GPD of the form:

$$F_Y(y) = \begin{cases} 1 - \left(1 + \gamma \frac{y}{\theta}\right)^{-1/\gamma}, & y > 0,\ \gamma \neq 0 \\ 1 - e^{-\gamma/\theta}, & y > 0,\ \gamma = 0. \end{cases} \quad (18)$$

When $\gamma < 0$, the highest age at death of individuals in the population is

$$\omega = N - \frac{\theta}{\gamma} = N + \frac{\theta}{|\gamma|}. \quad (19)$$

Results

Table 1: Estimated gamma, highest attained age ω, and threshold age N, female 1901 cohort

Regime	$\hat{\gamma}$	SE($\hat{\gamma}$)	$\hat{\omega}$	SE($\hat{\omega}$)	N
TLT, HMD only	-0.307	0.0143	106.33	0.526	91
STLT, HMD only	-0.237	0.0451	108.98	2.245	97
TLT, HMD+CBS	-0.240	0.0070	109.59	0.388	91
STLT, HMD+CBS	-0.191	0.0132	111.78	0.976	97

TLT v.s. STLT – HMD Only

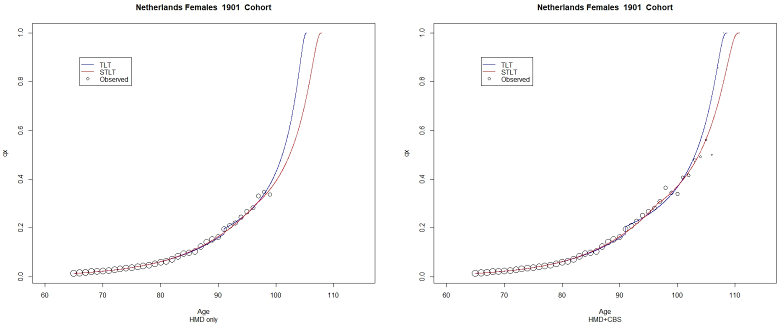


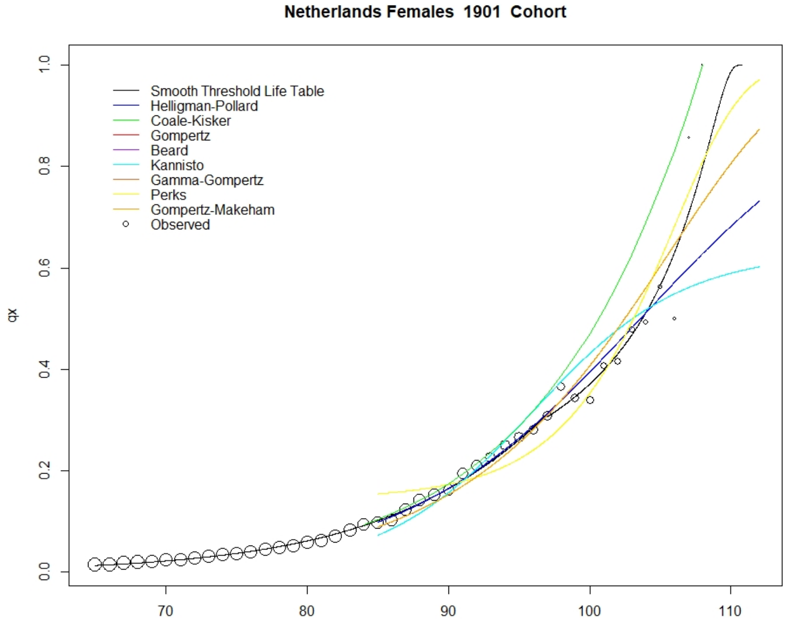
Figure 3: LEFT: HMD data; RIGHT: HMD+CBS.

TLT VS STLT – Goodness of fit

Table 2: SSEs under the TLT and STLT, HMD+CBS data

	Females		Males	
Cohort	TLT	STLT	TLT	STLT
1893	0.17	0.18	0.45	0.45
1894	0.20	0.16	0.48	0.50
1895	1.00	0.93	0.26	0.26
1896	0.16	0.10	0.61	0.69
1897	0.72	0.80	0.34	0.39
1898	0.10	0.08	0.35	0.28
1899	0.33	0.27	0.12	0.14
1900	0.84	0.74	0.54	0.50
1901	0.08	0.08	0.51	0.51

STLT v.s. Heligman-Pollard and Coale-Kisker



The fitted STLT for All Cohorts below Age 90

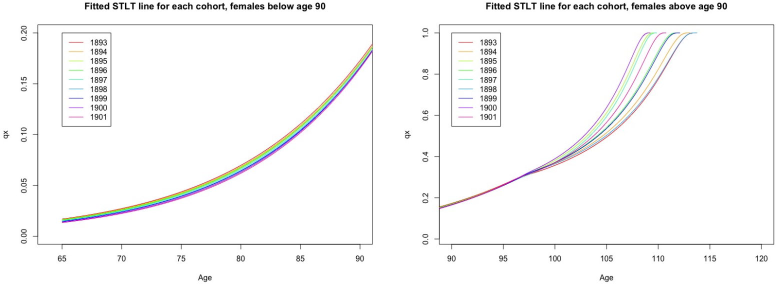


Figure 5: LEFT: Below age 90; RIGHT: Above age 90

The Dynamics of Fitted Parameters – Females

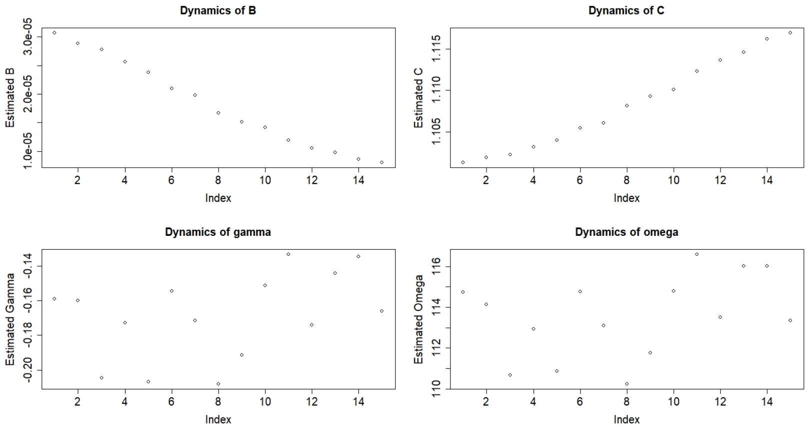


Figure 6: *The index corresponds to the cohort, eg. index 1 corresponds to the 1893 cohort.*

The Dynamic STLT Model

We model B in the STLT as the following function of time:

$$B_t = \exp(a + bt), \tag{20}$$

while keeping γ and $\theta = 1/(B_t C_t^N)$ constant through time. Then

$$C_t = (1/(\theta B_t))^{1/N} = \theta^{-N} \exp(-N(a + bt)) \tag{21}$$

N is also assumed to be constant across all cohorts, which is consistent with the compensation law of mortality (late-life mortality convergence), (e.g. Gavrilov & Gavrilova 1979, Gavrilov & Gavrilova 1991).

DSTLT–Model Fitting, Females

Fitting this model using MLE, we obtain the parameter estimates with the estimated threhold age $N = 98$.

Table 3: Estimated parameters for the DSTLT, females

Parameter	Estimate	Standard Error	95% Confidence Interval
a	-10.26	0.023	(-10.31, -10.22)
b	-0.085	0.0033	(-0.092, -0.079)
θ	2.58	0.0087	(2.56, 2.60)
γ	-0.174	0.0057	(-0.185, -0.163)

DSTLT v.s. CBD – Females

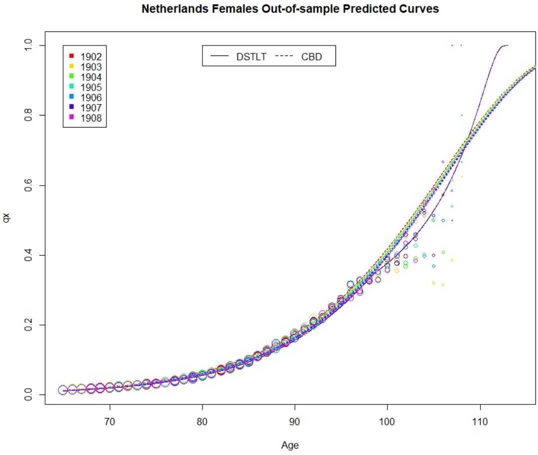


Figure 7: *Out-of-sample predicted DSTLT and CBD curves for the 1902-1908 female cohorts*

DSTLT v.s. CBD – Residuals, Females

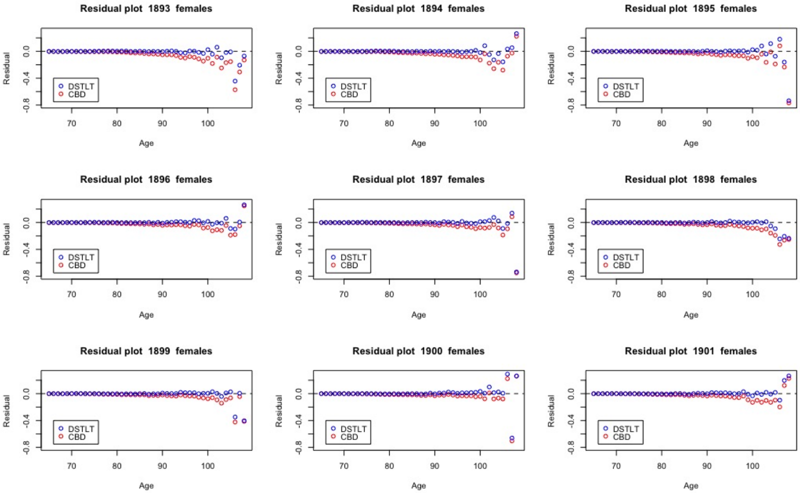


Figure 8: *DSTLT and CBD model residuals, females. Blue points correspond to the DSTLT; red points correspond to the CBD model*

Discussion 1– Tail Behavior of the Lifetime Distribution

1. The probability of dying within 1 year may reach 1 at a finite age ⇒ There is a fixed upper limit to the length of human life.
2. The probability remains below 1 at finite ages, but nevertheless tends asymptotically to 1. ⇒ "Life is unlimited but short".
3. The probability of dying within 1 year may asymptote to a limit less than 1.⇒ There is no fixed upper limit to the length of human life.

Discussion 2– Laws of Mortality

1. Gompertz-Makeham law of mortality
2. Compensation law of mortality
3. Late-life mortality deceleration

Our models satisfy all the three laws above plus a new law – "*advanced-age mortality acceleration*", which allows for a finite limit to human life span.